

STIC Biotechnology Systems Branch

RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/551,508
Source: PCT
Date Processed by STIC: 10/14/2005

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 4.2.2 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker/chkrnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry, Federal Express, United Parcel Service, or other delivery service (EFFECTIVE 01/14/05):
U.S. Patent and Trademark Office, Mail Stop Sequence, Customer Window, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

Revised 01/24/05

Raw Sequence Listing Error Summary

ERROR DETECTED

SUGGESTED CORRECTION

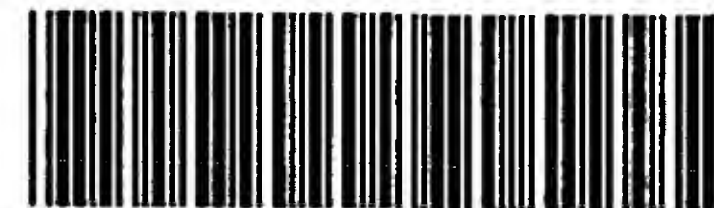
SERIAL NUMBER

10/551,508

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 **Wrapped Nucleics** The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 **Invalid Line Length** The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 **Misaligned Amino Numbering** The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 **Non-ASCII** The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 **Variable Length** Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 **PatentIn 2.0 "bug"** A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 **Skipped Sequences (OLD RULES)** Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 **Skipped Sequences (NEW RULES)** Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
(000)
- 9 **Use of n's or Xaa's (NEW RULES)** Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 **Invalid <213> Response** Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 **Use of <220>** Sequence(s) missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 **PatentIn 2.0 "bug"** Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 **Misuse of n/Xaa** "n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid



PCT

RAW SEQUENCE LISTING

DATE: 10/14/2005

PATENT APPLICATION: US/10/551,508

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

3 <110> APPLICANT: KIM, Young Tae
 4 LEE, Jae Hyung
 7 <120> TITLE OF INVENTION: Gene involved in the biosynthesis of carotenoid and marine
 8 microorganism, paracoccus haeundaesis, producing the
 9 carotenoid
 11 <130> FILE REFERENCE: 428.1056
 C--> 13 <140> CURRENT APPLICATION NUMBER: US/10/551,508
 C--> 13 <141> CURRENT FILING DATE: 2005-09-29
 13 <150> PRIOR APPLICATION NUMBER: PCT/KR2004/000752
 14 <151> PRIOR FILING DATE: 2003-03-31
 16 <150> PRIOR APPLICATION NUMBER: KR2003-20222
 17 <151> PRIOR FILING DATE: 2003-03-31
 19 <150> PRIOR APPLICATION NUMBER: KR2003-20023
 20 <151> PRIOR FILING DATE: 2003-03-31
 22 <160> NUMBER OF SEQ ID NOS: 18
 24 <170> SOFTWARE: KopatentIn 1.71

ERRORED SEQUENCES

52 <210> SEQ ID NO: 3
 53 <211> LENGTH: 1454
 54 <212> TYPE: DNA
 55 <213> ORGANISM: Paracoccus haeundaesis
 57 <400> SEQUENCE: 3
 58 caacttgaga gtttgatcct ggctcagaac gaacgctggc ggcaggctta acacatgcaa 60
 60 gtcgagcgag accttcgggt ctagcggcgg acgggtgagt aacgcgtggg aacgtgccct 120
 62 tctctacgga atagccccgg gaaactggga gtaataccgt atacgccctt tgggggaaag 180
 64 atttatcgga gaaggatcgg cccgcgttgg attaggtagt tgggtgggta atggcccacc 240
 66 aagccgacga tccatagctg gtttgagagg atgatcagcc acactgggac tgagacacgg 300
 68 cccagactcc tacgggaggg agcagtgggg aatcttagac aatgggggca accctgatct 360
 70 agccatgccg cgtgagtgat gaaggcctta ggggtgtaaa gctctttcag ctgggaagat 420
 72 aatgacggta ccagcagaag aagccccggc taactccgtg ccagcagccg cggtataacg 480
 74 gaggggggcta gcgttggtcg gaattactgg gcgtaaagcg cacgtgggcg gactggaaag 540
 76 tcagaggtga aatcccaggg ctcaaccttg gaactgcctt tgaaactatc agtctggagt 600
 78 tcgagagagg tgagtggaat tccgagtgtg gaggtgaaat tcgtagatat tcggaggaac 660
 80 accagtggcg aaggcggctc actggctcga tactgacgct gaggtgcgaa agcgtgggga 720
 82 gcaaacagga ttagataccc tggtagtcca cgccgtaaac gatgaatgcc agacgtcggc 780
 84 aagcatgctt gtcggtgtca cacctaacgg attaagcatt ccgcctgggg agtacggtcg 840
 86 caagattaaa actcaaagga attgacgggg gcccgcaaca gcggtggagc atgtggttta 900
 88 attcgaagca acgcgcagaa ccttaccac ccttgacatg gcaggaccgc tggagagatt 960
 90 cagctttctc gtaagagacc tgcacacagg tgctgcatgg ctgtcgtcag ctgctgtcgt 1020
 92 gagatgttcg gttaagtccg gcaacgagcg caaccacgt ccctagttgc cagcattcag 1080

Does Not Comply
 Corrected Diskette Needed
 (pg 1-9)

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

```

94 ttgggcactc tatggaaact gccgatgata agtcggagga aggtgtggat gacgtcaagt 1140
96 tctcatggcc cttacgggtt gggctacaca cgtgctacaa tgggtggtgac agtgggttaa 1200
98 tccccaaaag ccatctcagt tcggattgtc ctctgcaact cgagggcatg aagttggaat 1260
100 cgctagtaat cgcggaacag catgccgcgg tgaatacgtt cccgggcctt gtacacaccg 1320
102 cccgtcacac catgggagtt ggttctaccc gacgacgtg cgtaacctt cggggggcag 1380
104 gcggccacgg taggatcagc gactgggggtg aagtcgtaac aaggtagccg taggggaacc 1440
E--> 106 tgcggctgga tcac 1454
107 1454
325 <210> SEQ ID NO: 5
326 <211> LENGTH: 729
327 <212> TYPE: DNA
328 <213> ORGANISM: crtW gene
330 <400> SEQUENCE: 5
331 atgagcgcac atgccctgcc caaggcagat ctgaccgcca ccagcctgat cgtctcgggc 60
333 ggcacatcatg ccgcgtggct ggccctgcat gtgcatgcgc tgtggtttct ggacgcggcg 120
335 gcgcatccca tcctggcgat cgcgaatttc ctggggctga cctggctgtc ggtcgggtctg 180
337 ttcttcatcg cgcacatgac gatgcacggg tcggtcgtgc cggggcgctc gcgcggcaat 240
339 gcggcgatgg gccagctggg cctgtggctg tatgccggat ttctgtggcg caagatgatc 300
341 gtcaagcaca tggcccatca ccgccatacc ggaaccgacg acgaccccca ttctgaccat 360
343 ggccggcccg tccgctggta cgcgcgcttc atcggcacct atttcggctg gcgcgagggg 420
345 ctgctgctgc ccgtcatcgt gacggcttat gcgctgatcc tgggggatcg ctggatgtac 480
347 gtggtcttct ggccgctgcc gtcgatcctg gcgtcgatcc agctgttcgt gttcggcacc 540
349 tggttgccgc accgccccgg ccacgacgcg ttcccggacc gccataatgc gcggtcgtcg 600
351 cggatcagcg accccgtgtc gctgctgacc tgccttctact ttggtggtta tcatcacgaa 660
353 caccacctgc acccgacggt gccttggtgg cgcccgcca gcacccgcac caaggggggac 720
E--> 355 accgcatga 729
356 729
359 <210> SEQ ID NO: 6
360 <211> LENGTH: 242 OK
361 <212> TYPE: PRT
362 <213> ORGANISM: crtW amino acid
364 <400> SEQUENCE: 6
365 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu
E--> 366 1 5 10 10 15
368 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
E--> 369 20 20 25 25 30 30
371 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala
E--> 372 35 35 40 40 45 45
374 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Phe Ile Ala
E--> 375 50 55 55 60 60
377 His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Gly Asn
E--> 378 65 70 70 75 75 80
380 Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp
E--> 381 85 85 90 90 95 95
383 Arg Lys Met Ile Val Lys His Met Ala His His Arg His Thr Gly Thr
E--> 384 100 100 105 105 110 110
386 Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala
E--> 387 115 120 120 125 125
389 Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro
130 135 140

```

Mis-aligned
 Amino
 Acid No.
 Pls see
 Item #
 3
 on error
 Summary
 Sheet.

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

```

E--> 390      130      135      140
      392 Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
E--> 393 (45) (150) (150) (155) (155) (160) (160)
      395 Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe
E--> 396 (165) (165) (170) (170) (175) (175)
      398 Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro
E--> 399 (180) (180) (185) (185) (190) (190)
      401 Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu
E--> 402 (195) (200) (200) (205) (205)
      404 Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His
E--> 405 (210) (215) (220) (220)
      407 Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp
E--> 408 (225) (230) (235) (235) (240) (240)
      410 Thr Ala
      414 <210> SEQ ID NO: 7
      415 <211> LENGTH: 489
      416 <212> TYPE: DNA
      417 <213> ORGANISM: crtZ gene
      419 <400> SEQUENCE: 7
      420 atgaccaatt tcctgatcgt cgtcgccacc gtgctggtga tggagttgac ggcctattcc 60
      422 gtccaccggtt ggatcatgca cggccccctg ggctggggct ggcacaagtc ccaccacgag 120
      424 gaacacgacc acgcgctgga aaagaacgac ctgtacggcc tggctctttgc ggtgatcgcc 180
      426 acggtgctgt tcacgggtggg ctggatctgg gcgcccgtcc tgtggtggat cgctttgggc 240
      428 atgaccgtct atgggctgat ctatttcgtc ctgcatgacg ggctggttca tcagcgttg 300
      430 ccgttccgct atatcccgcg caagggctat gcccgccgcc tgtatcaggc ccaccgcctg 360
      432 caccacgcgg tcgagggacg cgaccattgc gtcagcttcg gcttcattca tgcgccgccg 420
      434 gtcgacaagc tgaagcagga cctgaagacg tcgggctgc tgcgggccga ggcgaggag 480
E--> 436 cgcacgtga 489
      437 489
      440 <210> SEQ ID NO: 8
      441 <211> LENGTH: 162
      442 <212> TYPE: PRT
      443 <213> ORGANISM: crtZ amino acid
      445 <400> SEQUENCE: 8
      446 Met Thr Asn Phe Leu Ile Val Val Ala Thr Val Leu Val Met Glu Leu
E--> 447 1 5 10 15
      449 Thr Ala Tyr Ser Val His Arg Trp Ile Met His Gly Pro Leu Gly Trp
E--> 450 20 25 30
      452 Gly Trp His Lys Ser His His Glu Glu His Asp His Ala Leu Glu Lys
E--> 453 35 40 45
      455 Asn Asp Leu Tyr Gly Leu Val Phe Ala Val Ile Ala Thr Val Leu Phe
E--> 456 50 55 60
      458 Thr Val Gly Trp Ile Trp Ala Pro Val Leu Trp Trp Ile Ala Leu Gly
E--> 459 65 70 75 80
      461 Met Thr Val Tyr Gly Leu Ile Tyr Phe Val Leu His Asp Gly Leu Val
E--> 462 85 90 95
      464 His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr Ala Arg
E--> 465 100 105 110
      467 Arg Leu Tyr Gln Ala His Arg Leu His His Ala Val Glu Gly Arg Asp

```

Same
Error

RAW SEQUENCE LISTING

DATE: 10/14/2005

PATENT APPLICATION: US/10/551,508

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

```

E--> 468      115      120      125
      470 His Cys Val Ser Phe Gly Phe Ile Tyr Ala Pro Pro Val Asp Lys Leu
E--> 471      130      135      140
      473 Lys Gln Asp Leu Lys Thr Ser Gly Val Leu Arg Ala Glu Ala Gln Glu
E--> 474 145      150      155      160
      476 Arg Thr
      527 <210> SEQ ID NO: 10
      528 <211> LENGTH: 386
      529 <212> TYPE: PRT
      530 <213> ORGANISM: crtY amino acid
      532 <400> SEQUENCE: 10
      533 Val Thr His Asp Val Leu Leu Ala Gly Ala Gly Leu Ala Asn Gly Leu
E--> 534 1 5 10 15
      536 Ile Ala Leu Ala Leu Arg Ala Ala Arg Pro Asp Leu Arg Val Leu Leu
E--> 537 20 25 30
      539 Leu Asp His Ala Ala Gly Pro Ser Asp Gly His Thr Trp Ser Cys His
E--> 540 35 40 45
      542 Asp Pro Asp Leu Ser Pro His Trp Leu Ala Arg Leu Lys Pro Leu Arg
E--> 543 50 55 60
      545 Arg Ala Asn Trp Pro Asp Gln Glu Val Arg Phe Pro Arg His Ala Arg
E--> 546 65 70 75 80
      548 Arg Leu Ala Thr Gly Tyr Gly Ser Leu Asp Gly Ala Ala Leu Ala Asp
E--> 549 85 90 95
      551 Ala Val Ala Arg Ser Gly Ala Glu Ile Arg Trp Asn Ser Asp Ile Ala
E--> 552 100 105 110
      554 Leu Leu Asp Glu Gln Gly Ala Thr Leu Ser Cys Gly Thr Arg Ile Glu
E--> 555 115 120 125
      557 Ala Gly Ala Val Leu Asp Gly Arg Gly Ala Gln Pro Ser Arg His Leu
E--> 558 130 135 140
      560 Thr Val Gly Phe Gln Lys Phe Val Gly Val Glu Ile Glu Thr Asp Cys
E--> 561 145 150 155 160
      563 Pro His Gly Val Pro Arg Pro Met Ile Met Asp Ala Thr Val Thr Gln
E--> 564 165 170 175
      566 Gln Asp Gly Tyr Arg Phe Ile Tyr Leu Leu Pro Phe Ser Pro Thr Arg
E--> 567 180 185 190
      569 Ile Leu Ile Glu Asp Thr Arg Tyr Ser Asp Gly Gly Asn Leu Asp Asp
E--> 570 195 200 205
      572 Asp Ala Leu Ala Ala Ala Ser His Asp Tyr Ala Arg Gln Gln Gly Trp
E--> 573 210 215 220
      575 Thr Gly Ala Glu Val Arg Arg Glu Arg Gly Ile Leu Pro Ile Ala Leu
E--> 576 225 230 235 240
      578 Ala His Asp Ala Ala Gly Phe Trp Ala Asp His Ala Glu Gly Pro Val
E--> 579 245 250 255
      581 Pro Val Gly Leu Arg Ala Gly Phe Phe His Pro Val Thr Gly Tyr Ser
E--> 582 260 265 270
      584 Leu Pro Tyr Ala Ala Gln Val Ala Asp Val Val Ala Gly Leu Ser Gly
E--> 585 275 280 285
      587 Pro Pro Gly Thr Asp Ala Leu Arg Gly Ala Ile Arg Asp Tyr Ala Ile
E--> 588 290 295 300

```

*Same
Error
Mis-aligned*

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

```

590 Asp Arg Ala Arg Arg Asp Arg Phe Leu Arg Leu Leu Asn Arg Met Leu
E--> 591 305          310          315          320
593 Phe Arg Gly Cys Ala Pro Asp Arg Arg Tyr Thr Leu Leu Gln Arg Phe
E--> 594          325          330          335
596 Tyr Arg Met Pro His Gly Leu Ile Glu Arg Phe Tyr Ala Gly Arg Leu
E--> 597          340          345          350
599 Ser Val Ala Asp Gln Leu Arg Ile Val Thr Gly Lys Pro Pro Ile Pro
E--> 600          355          360          365
602 Leu Gly Thr Ala Ile Arg Cys Leu Pro Glu Arg Pro Leu Leu Lys Glu
E--> 603          370          375          380

```

605 Asn Ala

606 385

609 <210> SEQ ID NO: 11

610 <211> LENGTH: 1506

611 <212> TYPE: DNA

612 <213> ORGANISM: crtI gene

614 <400> SEQUENCE: 11

```

615 atgaacgccc attcgcccgc ggccaagacc gccatcgtga tcggcgcagg ctttggcggg 60
617 ctggcccttg ccattccgct gcagtcgcgc ggcattcgcca ccaccttggg cgaggcccgg 120
619 gacaagcccg gcgggcgcgc ctatgtcttg cacgatcagg gccatgtctt cgacgcgggc 180
621 ccgaccgtca tcaccgaccc cgatgcgctc aaggagctgt gggcgctgac cgggcaggac 240
623 atggcgcgcg acgtgacgct gatgccggtg tcgcccttct atcgactgat gtggccgggc 300
625 gggaaggtct tcgattacgt gaacgaggcc gatcagctgg agcgccagat cgcccagttc 360
627 aaccgggacg acctggaagg ataccgccgc ttccgtgatt acgcggagga ggtgtatcag 420
629 gagggctacg tcaagctggg caccgtgccc ttccctcaagc tgggcccagat gctcaaggcc 480
631 gcgcccgcgc tgatgaagct ggaggcctat aagtcctgcc atgccaaggt cgcgaccttc 540
633 atcaaggacc cctatctgcg gcaggcggtt tcgtatcaca cgctgctggg gggcggggat 600
635 cccttctcga ccagctcgat ctatgcgctg atccacgcgc tggagcggcg cggcggggtc 660
637 tggttcgcca agggcggcac caaccagctg gtcgcgggca tggtcgcgct gttcgaacgg 720
639 cttggcggcc agatgatgct gaacgccaaag gtcgcccgga tcgagaccga gggcgcgcg 780
641 accacgggcg tcaccttggc ggacggggcg tctttaaggg ccgacatggg cgccagcaac 840
643 ggcgacgtca tgcacaacta tcgcgacctg ctgggcccaca cggcccgcgg gcagagccgc 900
645 gcgaaatcgc tggaccgcaa gcgctgggtcc atgtcgttgt tcgtgctgca ttctgggtctg 960
647 cgcgaggcgc ccaaggacat cgcgcatcac accatcctgt tcggcccccg ctacagggag 1020
649 ctggtcaacg agatcttcaa gggcccgaag ctggccgagg atttctcgt gtacctgcat 1080
651 tcgccctgca cgaccgatcc ggacatggcg cctccgggca tgtccacgca ttacgtgctg 1140
653 gcccccgctg cgcattctgg ccgcgccgag atcgattggg cggtcgaggg gccgcgctat 1200
655 gccgaccgca tcctggcgct cctggaggag cggctgatcc cgaacctgcg cgccaacctg 1260
657 accacgacgc gcattctcac gcccgccgat ttcgccagcg aactgaacgc ccatcacggc 1320
659 agcgcccttct cggtcgagcc gatcctgacg caatccgcgt ggttccggcc gcacaaccgc 1380
661 gacaagacga tccgcaactt ctatctggtc ggcgcgggca cccatccggg cgcgggcatt 1440
663 ccgggcgctc tgggctcggc caaggccacg gccaggtga tgctgtccga cctggcgggc 1500

```

E--> 665 gcatga

666 1506

669 <210> SEQ ID NO: 12

670 <211> LENGTH: 501

671 <212> TYPE: PRT

672 <213> ORGANISM: crtI amino acid

674 <400> SEQUENCE: 12

71506

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

	675	Met	Asn	Ala	His	Ser	Pro	Ala	Ala	Lys	Thr	Ala	Ile	Val	Ile	Gly	Ala		
E-->	676	1					5						10					15	
	678	Gly	Phe	Gly	Gly	Leu	Ala	Leu	Ala	Ile	Arg	Leu	Gln	Ser	Ala	Gly	Ile		
E-->	679				20					25							30		
	681	Ala	Thr	Thr	Leu	Val	Glu	Ala	Arg	Asp	Lys	Pro	Gly	Gly	Arg	Ala	Tyr		
E-->	682				35				40						45				
	684	Val	Trp	His	Asp	Gln	Gly	His	Val	Phe	Asp	Ala	Gly	Pro	Thr	Val	Ile		
E-->	685		50						55						60				
	687	Thr	Asp	Pro	Asp	Ala	Leu	Lys	Glu	Leu	Trp	Ala	Leu	Thr	Gly	Gln	Asp		
E-->	688	65					70						75				80		
	690	Met	Ala	Arg	Asp	Val	Thr	Leu	Met	Pro	Val	Ser	Pro	Phe	Tyr	Arg	Leu		
E-->	691						85					90					95		
	693	Met	Trp	Pro	Gly	Gly	Lys	Val	Phe	Asp	Tyr	Val	Asn	Glu	Ala	Asp	Gln		
E-->	694				100					105						110			
	696	Leu	Glu	Arg	Gln	Ile	Ala	Gln	Phe	Asn	Pro	Asp	Asp	Leu	Glu	Gly	Tyr		
E-->	697				115					120						125			
	699	Arg	Arg	Phe	Arg	Asp	Tyr	Ala	Glu	Glu	Val	Tyr	Gln	Glu	Gly	Tyr	Val		
E-->	700		130						135						140				
	702	Lys	Leu	Gly	Thr	Val	Pro	Phe	Leu	Lys	Leu	Gly	Gln	Met	Leu	Lys	Ala		
E-->	703	145					150						155				160		
	705	Ala	Pro	Ala	Leu	Met	Lys	Leu	Glu	Ala	Tyr	Lys	Ser	Val	His	Ala	Lys		
E-->	706					165					170						175		
	708	Val	Ala	Thr	Phe	Ile	Lys	Asp	Pro	Tyr	Leu	Arg	Gln	Ala	Phe	Ser	Tyr		
E-->	709				180						185					190			
	711	His	Thr	Leu	Leu	Val	Gly	Gly	Asn	Pro	Phe	Ser	Thr	Ser	Ser	Ile	Tyr		
E-->	712				195					200						205			
	714	Ala	Leu	Ile	His	Ala	Leu	Glu	Arg	Arg	Gly	Gly	Val	Trp	Phe	Ala	Lys		
E-->	715		210						215					220					
	717	Gly	Gly	Thr	Asn	Gln	Leu	Val	Ala	Gly	Met	Val	Ala	Leu	Phe	Glu	Arg		
E-->	718	225				230						235					240		
	720	Leu	Gly	Gly	Gln	Met	Met	Leu	Asn	Ala	Lys	Val	Ala	Arg	Ile	Glu	Thr		
E-->	721					245					250						255		
	723	Glu	Gly	Ala	Arg	Thr	Thr	Gly	Val	Thr	Leu	Ala	Asp	Gly	Arg	Ser	Leu		
E-->	724				260						265					270			
	726	Arg	Ala	Asp	Met	Val	Ala	Ser	Asn	Gly	Asp	Val	Met	His	Asn	Tyr	Arg		
E-->	727			275						280					285				
	729	Asp	Leu	Leu	Gly	His	Thr	Ala	Arg	Gly	Gln	Ser	Arg	Ala	Lys	Ser	Leu		
E-->	730		290						295					300					
	732	Asp	Arg	Lys	Arg	Trp	Ser	Met	Ser	Leu	Phe	Val	Leu	His	Phe	Gly	Leu		
E-->	733	305				310						315					320		
	735	Arg	Glu	Ala	Pro	Lys	Asp	Ile	Ala	His	His	Thr	Ile	Leu	Phe	Gly	Pro		
E-->	736					325					330						335		
	738	Arg	Tyr	Arg	Glu	Leu	Val	Asn	Glu	Ile	Phe	Lys	Gly	Pro	Lys	Leu	Ala		
E-->	739				340						345					350			
	741	Glu	Asp	Phe	Ser	Leu	Tyr	Leu	His	Ser	Pro	Cys	Thr	Thr	Asp	Pro	Asp		
E-->	742			355					360						365				
	744	Met	Ala	Pro	Pro	Gly	Met	Ser	Thr	His	Tyr	Val	Leu	Ala	Pro	Val	Pro		
E-->	745		370					375					380						
	747	His	Leu	Gly	Arg	Ala	Glu	Ile	Asp	Trp	Ala	Val	Glu	Gly	Pro	Arg	Tyr		

*same
error*

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

```

E--> 748 385          390          395          400
      750 Ala Asp Arg Ile Leu Ala Ser Leu Glu Glu Arg Leu Ile Pro Asn Leu
E--> 751          405          410          415
      753 Arg Ala Asn Leu Thr Thr Thr Arg Ile Phe Thr Pro Ala Asp Phe Ala
E--> 754          420          425          430
      756 Ser Glu Leu Asn Ala His His Gly Ser Ala Phe Ser Val Glu Pro Ile
E--> 757          435          440          445
      759 Leu Thr Gln Ser Ala Trp Phe Arg Pro His Asn Arg Asp Lys Thr Ile
E--> 760          450          455          460
      762 Arg Asn Phe Tyr Leu Val Gly Ala Gly Thr His Pro Gly Ala Gly Ile
E--> 763 465          470          475          480
      765 Pro Gly Val Val Gly Ser Ala Lys Ala Thr Ala Gln Val Met Leu Ser
E--> 766          485          490          495
      768 Asp Leu Ala Gly Ala
E--> 769          500
      772 <210> SEQ ID NO: 13
      773 <211> LENGTH: 915
      774 <212> TYPE: DNA
      775 <213> ORGANISM: crtB gene
      777 <400> SEQUENCE: 13
      778 atgagcgcgc tggctcctgac ctgcgaccgag gcgatcaccc aagggtcgca aagctttgcc 60
      780 acggcgccca agctgatgcc gccgggcacg cgcgacgaca cggatgatgct ctatgcctgg 120
      782 tgccgccacg cggatgacgt gatcgacggt caggccctgg gcagccgccc cgaggcgggtg 180
      784 aacgaccgcg aggcgcggct ggacggcctg cgcgtcgaca cgctggcgcc cctgcagggc 240
      786 gacgggtccg tgaccccgcc ctttgccgcg ctgcgcgcgg tggcgcgccg gcatgatttc 300
      788 ccgcaggcct ggcccatgga cctgatcgaa ggcttcgcga tggatgtcga ggcgcgcgac 360
      790 tatcgcacgc tggatgacgt gctggaatat tctatcacg tcgcaggcat cgtcggcggtg 420
      792 atgatggccc gcgtgatggg cgtgcgcgac gatcctgtcc tggaccgcgc ctgcgacctg 480
      794 gggctggcgt tccagctgac caacatcgcg cgcgacgtga tcgacgatgc gcgcacggg 540
      796 cggatgctat tgccggggga ctggctggac caggcgggcg cgcggatcga cgggcccgggtg 600
      798 ccgtcgccgg agctgtacac agtgatcctc cggctgttgg atgaggcgga accctattac 660
      800 gcgtcgccgc ggggtgggtct ggcggtatct ccaccgcgct gcgcctggtc catcgccgcc 720
      802 gcgctacgga tctatcgcg catcgggctg cgcacccgca agagcgggcc gcaggcctat 780
      804 cgccagcgga tcagcacgtc caaggctgcc aagatcggcc tgctgggctg cgggggctgg 840
      806 gatgtcgcg gatcacgct gccgggggcg ggcgtgtcgc ggcagggcct ctggaccg 900
E--> 808 ccgcatcacg tctag
      809 915
      812 <210> SEQ ID NO: 14
      813 <211> LENGTH: 304
      814 <212> TYPE: PRT
      815 <213> ORGANISM: crtB amino acid
      817 <400> SEQUENCE: 14
      818 Met Ser Asp Leu Val Leu Thr Ser Thr Glu Ala Ile Thr Gln Gly Ser
E--> 819 1          5          10          15
      821 Gln Ser Phe Ala Thr Ala Ala Lys Leu Met Pro Pro Gly Ile Arg Asp
E--> 822          20          25          30
      824 Asp Thr Val Met Leu Tyr Ala Trp Cys Arg His Ala Asp Asp Val Ile
E--> 825          35          40          45
      827 Asp Gly Gln Ala Leu Gly Ser Arg Pro Glu Ala Val Asn Asp Pro Gln

```

Same
Error

7915

Same
Error

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

E--> 828 50 55 60
 830 Ala Arg Leu Asp Gly Leu Arg Val Asp Thr Leu Ala Ala Leu Gln Gly
 E--> 831 65 70 75 80
 833 Asp Gly Pro Val Thr Pro Pro Phe Ala Ala Leu Arg Ala Val Ala Arg
 E--> 834 85 90 95
 836 Arg His Asp Phe Pro Gln Ala Trp Pro Met Asp Leu Ile Glu Gly Phe
 E--> 837 100 105 110
 839 Ala Met Asp Val Glu Ala Arg Asp Tyr Arg Thr Leu Asp Asp Val Leu
 E--> 840 115 120 125
 842 Glu Tyr Ser Tyr His Val Ala Gly Ile Val Gly Val Met Met Ala Arg
 E--> 843 130 135 140
 845 Val Met Gly Val Arg Asp Asp Pro Val Leu Asp Arg Ala Cys Asp Leu
 E--> 846 145 150 155 160
 848 Gly Leu Ala Phe Gln Leu Thr Asn Ile Ala Arg Asp Val Ile Asp Asp
 E--> 849 165 170 175
 851 Ala Arg Ile Gly Arg Cys Tyr Leu Pro Gly Asp Trp Leu Asp Gln Ala
 E--> 852 180 185 190
 854 Gly Ala Arg Ile Asp Gly Pro Val Pro Ser Pro Glu Leu Tyr Thr Val
 E--> 855 195 200 205
 857 Ile Leu Arg Leu Leu Asp Glu Ala Glu Pro Tyr Tyr Ala Ser Ala Arg
 E--> 858 210 215 220
 860 Val Gly Leu Ala Asp Leu Pro Pro Arg Cys Ala Trp Ser Ile Ala Ala
 E--> 861 225 230 235 240
 863 Ala Leu Arg Ile Tyr Arg Ala Ile Gly Leu Arg Ile Arg Lys Ser Gly
 E--> 864 245 250 255
 866 Pro Gln Ala Tyr Arg Gln Arg Ile Ser Thr Ser Lys Ala Ala Lys Ile
 E--> 867 260 265 270
 869 Gly Leu Leu Gly Val Gly Gly Trp Asp Val Ala Arg Ser Arg Leu Pro
 E--> 870 275 280 285
 872 Gly Ala Gly Val Ser Arg Gln Gly Leu Trp Thr Arg Pro His His Val
 E--> 873 290 295 300
 915 <210> SEQ ID NO: 16
 916 <211> LENGTH: 293
 917 <212> TYPE: PRT
 918 <213> ORGANISM: crtE amino acid
 920 <400> SEQUENCE: 16
 921 Met Arg Arg Asp Val Asn Pro Ile His Ala Thr Leu Leu Gln Thr Arg
 E--> 922 1 5 10 15
 924 Leu Glu Glu Ile Ala Gln Gly Phe Gly Ala Val Ser Gln Pro Leu Gly
 E--> 925 20 25 30
 927 Ala Ala Met Ser His Gly Ala Leu Ser Ser Gly Arg Arg Phe Arg Gly
 E--> 928 35 40 45
 930 Met Leu Met Leu Leu Ala Ala Glu Ala Ser Gly Gly Val Cys Asp Thr
 E--> 931 50 55 60
 933 Ile Val Asp Ala Ala Cys Ala Val Glu Met Val His Ala Ala Ser Leu
 E--> 934 65 70 75 80
 936 Ile Phe Asp Asp Leu Pro Cys Met Asp Asp Ala Gly Leu Arg Arg Gly
 E--> 937 85 90 95
 939 Arg Pro Ala Thr His Val Ala His Gly Glu Ser Arg Ala Val Leu Gly

*Same
 2nd*

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005

TIME: 11:04:23

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

E--> 940	100	105	110
942 Gly Ile Ala Leu Ile Thr Glu Ala Met Ala Leu Leu Ala Gly Ala Arg			
E--> 943	115	120	125
945 Gly Ala Ser Gly Thr Val Arg Ala Gln Leu Val Arg Ile Leu Ser Arg			
E--> 946	130	135	140
948 Ser Leu Gly Pro Gln Gly Leu Cys Ala Gly Gln Asp Leu Asp Leu His			
E--> 949 145	150	155	160
951 Ala Ala Lys Asn Gly Ala Gly Val Glu Gln Glu Gln Asp Leu Lys Thr			
E--> 952	165	170	175
954 Gly Val Leu Phe Ile Ala Gly Leu Glu Met Leu Ala Val Ile Lys Glu			
E--> 955	180	185	190
957 Phe Asp Ala Glu Glu Gln Thr Gln Met Ile Asp Phe Gly Arg Gln Leu			
E--> 958	195	200	205
960 Gly Arg Val Phe Gln Ser Tyr Asp Asp Leu Leu Asp Val Val Gly Asp			
E--> 961	210	215	220
963 Gln Ala Ala Leu Gly Lys Asp Thr Gly Arg Asp Ala Ala Ala Pro Gly			
E--> 964 225	230	235	240
966 Pro Arg Arg Gly Leu Leu Ala Val Ser Asp Leu Gln Asn Val Ser Arg			
E--> 967	245	250	255
969 His Tyr Glu Ala Ser Arg Ala Gln Leu Asp Ala Met Leu Arg Ser Lys			
E--> 970	260	265	270
972 Arg Leu Gln Ala Pro Glu Ile Ala Ala Leu Leu Glu Arg Val Leu Pro			
E--> 973	275	280	285
975 Tyr Ala Ala Arg Ala			
976	290		

same
error

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005
TIME: 11:04:24

Input Set : N:\FANTU\10551508.raw.txt
Output Set: N:\CRF4\10142005\J551508.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 7,8

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/551,508

DATE: 10/14/2005

TIME: 11:04:24

Input Set : N:\FANTU\10551508.raw.txt

Output Set: N:\CRF4\10142005\J551508.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application No
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:106 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:1454 SEQ:3 ✓
L:355 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:729 SEQ:5 ✓
L:366 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:6 ✓
M:332 Repeated in SeqNo=6
L:436 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:489 SEQ:7 ✓
L:447 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:8 ✓
M:332 Repeated in SeqNo=8
L:534 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:10 ✓
M:332 Repeated in SeqNo=10
L:665 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:1506 SEQ:11 ✓
L:676 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:12 ✓
M:332 Repeated in SeqNo=12
L:808 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:915 SEQ:13 ✓
L:819 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:14 ✓
M:332 Repeated in SeqNo=14
L:922 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:16 ✓
M:332 Repeated in SeqNo=16